AFRL-RX-WP-TP-2011-4405

PROBABILISTIC COLLOCATION METHOD FOR NDE PROBLEMS WITH UNCERTAIN PARAMETERS WITH ARBITRARY DISTRIBUTIONS (PREPRINT)

Jeremy Knopp and Mark Blodgett

Metals, Ceramics & Nondestructive Evaluation Division (AFRL/RXLP)

M.R. Cherry

University of Dayton Research Institute

NOVEMBER 2011

Approved for public release; distribution unlimited.

See additional restrictions described on inside pages

STINFO COPY

AIR FORCE RESEARCH LABORATORY
MATERIALS AND MANUFACTURING DIRECTORATE
WRIGHT-PATTERSON AIR FORCE BASE, OH 45433-7750
AIR FORCE MATERIEL COMMAND
UNITED STATES AIR FORCE

REPORT DOCUMENTATION PAGE	*Form Approved* *OMB No. 0704-0188*

1. REPORT DATE *(DD-MM-YY)*	2. REPORT TYPE	3. DATES COVERED *(From - To)*
November 2011	Technical Paper	1 October 2011 – 1 October 2011

4. TITLE AND SUBTITLE
PROBABILISTIC COLLOCATION METHOD FOR NDE PROBLEMS WITH UNCERTAIN PARAMETERS WITH ARBITRARY DISTRIBUTIONS (PREPRINT)

5a. CONTRACT NUMBER In-house
5b. GRANT NUMBER
5c. PROGRAM ELEMENT NUMBER 62102F

6. AUTHOR(S)
Jeremy Knopp and Mark Blodgett (AFRL/RXLP)
M.R. Cherry (University of Dayton Research Institute)

5d. PROJECT NUMBER 4349
5e. TASK NUMBER 40
5f. WORK UNIT NUMBER LP111200

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Nondestructive Evaluation Branch/Metals, Ceramics & Nondestructive Evaluation Division
Air Force Research Laboratory, Materials and Manufacturing Directorate
Wright-Patterson Air Force Base, OH 45433-7750
Air Force Materiel Command, United States Air Force

University of Dayton Research Institute

8. PERFORMING ORGANIZATION REPORT NUMBER
AFRL-RX-WP-TP-2011-4405

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
Air Force Research Laboratory
Materials and Manufacturing Directorate
Wright-Patterson Air Force Base, OH 45433-7750
Air Force Materiel Command
United States Air Force

10. SPONSORING/MONITORING AGENCY ACRONYM(S)
AFRL/RXLP

11. SPONSORING/MONITORING AGENCY REPORT NUMBER(S)
AFRL-RX-WP-TP-2011-4405

12. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release; distribution unlimited.

13. SUPPLEMENTARY NOTES
The U.S. Government is joint author of this work and has the right to use, modify, reproduce, release, perform, display, or disclose the work. PA Case Number and clearance date: 88ABW-2011-5440, 12 Oct 2011. Preprint journal article to be submitted to Quantitative Nondestructive Evaluation 2011. This document contains color.

14. ABSTRACT
In order to quantify the reliability of NDE systems, large amounts of experiments are performed to develop a probability of detection (POD) curve for the system. These POD studies require a substantial amount of experimentation which can sometimes be cost prohibitive. To expedite the process of developing these curves, highly precise numerical models are used in conjunction with NDE sensors to understand the uncertainties associated with the inspections. Numerical models are also used in stochastic inversion methods such as Bayesian inversion, which provide a means of characterizing system properties with uncertainties. A strong basis has been developed in the modeling and simulation community for deterministic forward models in NDE, but to fully incorporate these models in model-assisted probability of detection (MAPOD) studies or stochastic inversion schemes, the models must be treated in a stochastic sense. A method of taking random inputs to a "black box" forward model and developing the full probability distribution function (PDF) of the response has been proposed.

15. SUBJECT TERMS
probabilistic collocation method, uncertainty propogation

16. SECURITY CLASSIFICATION OF: a. REPORT	b. ABSTRACT	c. THIS PAGE	17. LIMITATION OF ABSTRACT:	NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON (Monitor)
Unclassified	Unclassified	Unclassified	SAR	10	Mark Blodgett
					19b. TELEPHONE NUMBER *(Include Area Code)* N/A

PROBABILISTIC COLLOCATION METHOD FOR NDE PROBLEMS WITH UNCERTAIN PARAMETERS WITH ARBITRARY DISTRIBUTIONS

M. R. Cherry[2], J. S. Knopp[1], M. P. Blodgett[1]

[1]Metals, Ceramics and NDE Division, AFRL, WPAFB, Dayton, OH, USA
[2]University of Dayton Research Institute, Structural Integrity Division, Dayton, OH, USA

ABSTRACT. In order to quantify the reliability of NDE systems, large amounts of experiments are performed to develop a probability of detection (POD) curve for the system. These POD studies require a substantial amount of experimentation which can sometimes be cost prohibitive. To expedite the process of developing these curves, highly precise numerical models are used in conjunction with NDE sensors to understand the uncertainties associated with the inspections. Numerical models are also used in stochastic inversion methods such as Bayesian inversion, which provide a means of characterizing system properties with uncertainties. A strong basis has been developed in the modeling and simulation community for deterministic forward models in NDE, but to fully incorporate these models in model-assisted probability of detection (MAPOD) studies or stochastic inversion schemes, the models must be treated in a stochastic sense. A method of taking random inputs to a "black box" forward model and developing the full probability distribution function (PDF) of the response has been proposed. This method, called the probabilistic collocation method (PCM), takes random inputs to a forward model and uses orthogonal polynomials to construct a surrogate model in the area of the expected values of the inputs which is solved much quicker than the original forward model. In the NDE community, this method has only been used with inputs of known, named distributions. In this work, inputs of arbitrary distribution were used and the orthogonal polynomials for these inputs were developed with a recursion relationship that has been shown to produce orthogonal polynomials with respect to a given, continuous function. A concise code was written to make testing the method and incorporating it into MAPOD studies and inversion schemes relatively easy. The routine was tested with academic problems as well as eddy current problems.

INTRODUCTION

Physics based computational models have been developed for a wide variety of NDE methodologies using many different numerical methods [e.g. 1,2,3]. Many of these models have been further developed into commercially available software packages such as Vic3D, ECSIM, PZFLEX, and CIVA, among others. Typically, these forward models are deterministic which implies the requirement of exact knowledge of problem specifics, such as material parameters, geometry parameters, and boundary conditions. In practice however, these parameters are rarely represented accurately by a single value. Rather, the exact value of the parameter is uncertain, and the parameter must be represented in the forward model as a random variable. This causes problems in forward models because the value of the parameter could then take on any value within the range of interest. The full

probability distribution of the output of the forward model is needed for inversion methods and MAPOD studies using a Bayesian framework. There have been many methods proposed to accomplish this task, many of which can be found in [4]. The most straightforward approach, called the Monte Carlo sampling method, is to sample the uncertain parameters at an adequate amount of points and solve the forward model at these points. The values can then be used to generate moment information about the distribution of the output or to construct histograms of the output values to determine the overall shape of the probability distribution function (PDF). While advancements have been made in reducing the sampling points needed to accurately represent these parameters of the output distribution, MC methods still require many forward model runs and can be costly. Another method to solve stochastic problems of this type is called the stochastic spectral finite element method [5]. In this method, the uncertain parameters are expanded in terms of their truncated Karhunen-Loève approximations which results in a system of equations to be solved with the finite element method. The issue with this method is that the problem requires development theoretically and numerically, whereas most commercially available software that solve NDE problems cannot be developed further. Ideally the UQ method of choice would be non-intrusive, i.e. would not require tampering with the program code that has already been developed. In 1938, Wiener proposed a new method to represent a stochastic process, called the Wiener process, that described the random motion (Brownian motion) of atoms in a drop of water [7]. This method involved representing the stochastic process as a series expansion of polynomials. In his study, the polynomials chosen (Hermite Polynomials) were orthogonal with respect to the PDF of a Gaussian distribution. From this, the polynomial chaos expansion (PCE) was defined and used to represent several different distributions. The PCE of the distributions can be used to represent functions of the distributions, which then leads directly to the PCM. In this work, the PCM is used to model uncertainties in a forward model for eddy current testing (ECT) in NDE. The PCM is extended to inputs with arbitrary distributions to account for the many different uncertain parameters in NDE applications.

ORTHOGONAL POLYNOMIALS

A gPC basis for a distribution is defined as a sequence of polynomials,

$$\varphi_i, i = 1, \dots, n$$

that are orthogonal with respect to the density function of the distribution [9]. To define orthogonal in this setting, let:

$$\left(\varphi_i(x), \varphi_j(x)\right)_{dm} = \int_{\Omega} \varphi_i(x)\varphi_j(x)dm(x) \tag{1}$$

be an inner product with respect to the measure $dm(x)$ where Ω is the support of x, or the range of possible values that x could take on. The measure is a generalization of the notion of length or size in Euclidean space, and is always positive. In probability theory, the measure is defined as:

$$p(x)dx \tag{2}$$

where $p(x)$ is the probability density function of the random variable. Functions that are orthogonal with respect to the measure, and thereby the PDF of the random variable satisfy the relation:

$$(u(x), v(x))_{p(x)} = \int_{\Omega} u(x)v(x)p(x)dx = 0 \tag{3}$$

and a sequence of polynomials are orthogonal if, for all $i, j = 1, \dots, n$:

$$\left(\varphi_i(x), \varphi_j(x)\right)_p = \begin{cases} \|\varphi_i(x)\|^2, & i = j \\ 0, & i \neq j \end{cases} \tag{4}$$

RECURRENCE RELATION

The approach above is valid if the polynomials orthogonal with respect to the given distribution are of known form. Several distributions and their respective orthogonal polynomials are documented in TABLE. For an arbitrarily distributed random variable, these polynomials may not be known. The recurrence relation in (5) has been shown [6,8] to produce orthogonal polynomials of increasing degree for a positive, Riemann integrable function $p(x)$.

$$\begin{aligned}
&\varphi_0(x) = 1 \\
&\varphi_1(x) = x - \frac{(x\varphi_0, \varphi_0)_p}{\|\varphi_0\|_p^2} \\
&\vdots \\
&\varphi_n = (x - A_n)\varphi_{n-1}(x) - B_n\varphi_{n-2}(x), \;\; n = 2,3, \dots \\
&A_n = \frac{(x\varphi_{n-1}, \varphi_{n-1})_p}{\|\varphi_{n-1}\|_p^2} \\
&B_n = \frac{(x\varphi_{n-1}, \varphi_{n-2})_p}{\|\varphi_{n-2}\|_p^2}
\end{aligned} \tag{5}$$

Since the calculation of A_n and B_n require the evaluation of the inner product and its norm, numerical integration of these values is needed. Fortunately, this integration is relatively straight forward as it is always in one dimension. In the case studies presented here, integration was performed with the trapezoidal method shown in [11].

STOCHASTIC PROBLEM

In this case study, the problem to be solved is change in impedance of an induction coil due to eddy currents induced in a conductive sample. Damage in the sample is not considered to make computations straightforward. In equation form, the deterministic problem can be written as:

$$Z = Z(\boldsymbol{r}) \tag{6}$$

where $\boldsymbol{r}$ is a vector with elements that are parameters of the problem, such as the sample conductivity, the liftoff of the probe, the length of the truncated domain on which the problem is defined, etc. The deterministic problem is well defined and many methods exist to solve it, but when the parameters of the problem are uncertain the problem needs to be reformulated. The new problem can be written as:

$$Z = Z(\boldsymbol{g}, \boldsymbol{r}') \tag{7}$$

where $\boldsymbol{g}$ is the vector of the random parameters and $\boldsymbol{r}'$ is the vector of the rest of the parameters. A problem with multiple uncertain parameters can be represented as a problem with one uncertain vector, and this vector has its own distribution and therefore, its own orthogonal polynomials. If the elements of the vector are independent, the distribution function of the vector, called the joint PDF, can be written as:

$$p_{\boldsymbol{g}} = \prod_{i=1}^{n} p_i \tag{8}$$

where n is the number of uncertain parameters and p_i is the PDF of the i^{th} distribution. The polynomials orthogonal to this random vector are defined as [9]:

$$\varphi_{\boldsymbol{I}} = \prod_{j=1}^{n} \varphi_{\boldsymbol{I}(j)}^{(j)} \tag{9}$$

where $\boldsymbol{I}$ is a n-dimensional multi-index and $\varphi_{\boldsymbol{I}(j)}^{(j)}$ is the polynomial of order $\boldsymbol{I}(j)$ that is orthogonal with respect to the j^{th} random variable. The sum of the components of $\boldsymbol{I}$ should not exceed the maximum order of polynomial desired for the polynomial fit for the problem, N. This will be clarified in a later section.

PROBABILISTIC COLLOCATION METHOD

Once the orthogonal polynomials for the random parameters vector are determined, the desired quantity can be represented as a linear combination of these polynomials [9]. In this case, we have:

$$Z(\boldsymbol{g}) = \sum_{i=1}^{l} a_i \varphi_{\boldsymbol{I}_i} \tag{10}$$

Here, l is the number of polynomials in the truncated sequence orthogonal to the random vector and a_i are the coefficients of each polynomial to be determined. Clearly, once these coefficients are determined, we have a polynomial surrogate representation of the actual model response, $Z(\boldsymbol{g})$. To find these coefficients, a set of points, called collocation points, is formed and the actual model is solved at these points. The orthogonal polynomials are then also evaluated at these points and the residual between the actual solution and the surrogate solution is set to zero, resulting in a linear system to be solved for the coefficient

values, $\boldsymbol{v}$, as shown in (11). This approach is analogous to basic polynomial interpolation, except the polynomial bases are defined differently.

$$A * \boldsymbol{v} = \boldsymbol{u} \tag{11}$$

$$\boldsymbol{u} = Z(\boldsymbol{g}_i)$$
$$A = [a_{i,j}] = \varphi_{\boldsymbol{I}_j}(\boldsymbol{g}_i)$$

Collocation points can be chosen by a number of methods. In this study, the roots of the higher order polynomials were chosen as collocation points. The roots of orthogonal polynomials are always real and distinct, and always lie in the interval of support of the distribution. The roots would also be used as Gauss points in a quadrature estimate of the moments of the distribution of impedance, and are thus the optimal points for this type of analysis. Other methods, such as sparse grid selection, can offer more information at the tails of the distributions and better convergence, but from experience, the roots provide adequate convergence for the purposes of eddy current problems with low dimensionality [].

CONVERGENCE MEASURES

To evaluate the validity of the surrogate model, the forward model must be sampled again and compared to the surrogate model by some means. For this study, these points were chosen as the 2nd higher order polynomial roots for the sake of computational savings in the event that the approximation is not adequate at the current polynomial order. The solutions of the forward model and the polynomial surrogate can be compared with the 2-norm:

$$\varepsilon = \left\| \boldsymbol{u}_m - \boldsymbol{u}_{app} \right\|_2 \tag{12}$$

Another method of comparison is the probabilistic sum of the square residuals:

$$ssr = \sqrt{\frac{\sum_{i=1}^{n} \epsilon_i}{n p_{\boldsymbol{g}}(\boldsymbol{g}_\mu)}} \tag{13}$$

$$\epsilon_i = \left(u_{app,i} - u_{m,i}\right)^2 p_{\boldsymbol{g}}(\boldsymbol{g}_i)$$

or the relative measure:

$$rssr = \frac{ssr}{E[u_{app}(\boldsymbol{g})]} \tag{14}$$

In these measures, $p_{\boldsymbol{g}}(\boldsymbol{g})$ is the joint probability distribution function of the random vector, r_μ is a vector of the means of the random parameters, $u_m(\boldsymbol{g})$ is the forward model and $u_{app}(\boldsymbol{g})$ is the surrogate model.

TABLE 1. A list of the distributions for each case study. The distributions in case study 2 were defined numerically to show that the results compare the results from case 1.

	Liftoff Distribution	Conductivity Distribution
Case Study 1	Uniform - U[1.53,2.53]	Normal – N[19.2,7.5]
Case Study 2	Uniform – U[1.53,2.53] (Def. Numerically)	Normal – N[19.2,7.5] (Def. Numerically)
Case Study 3	Triangular Tri[1.53,2.53,2.03]	Gamma

CASE STUDIES

In this study, the problem chosen for analysis is the well studied problem of the change of impedance of an induction coil when placed over a conducting half space. An analytical solution for this problem was developed by Dodd and Deeds [12]. Their integral solution requires numerical integration in its raw form, so the series solution of these integrals found in [10] was used in this study. Because the series solution can be solved relatively quickly, a full Monte Carlo simulation on the numerical model can be performed within reasonable time, and the results from the PCM development can be compared directly to the model results. The uncertain variables in this study were conductivity and liftoff because typical eddy current benchmark studies require inversion of this analytical model to determine these parameters. To accomplish this inversion in a probabilistic setting, a forward model that determines the distribution of impedance is crucial. The distributions chosen for the different studies are shown in Table 1. For comparison purposes, the parameters in the first case study were chosen to be distributed normally and uniformly, and the polynomials used were the associated Hermite and Legendre polynomials, respectively. In the second study, these distributions were defined numerically and their orthogonal polynomials were determined with numerical integration. The polynomials are compared to those from the first study, as well as the results from the simulation. In the third study, liftoff had a triangular distribution and conductivity had a gamma distribution. The triangular distribution has no associated orthogonal polynomials and the gamma polynomials were determined numerically.

RESULTS AND DISCUSSION

In all cases, the surrogate model converged at third order. The final PDF's from the surrogate model and the Dodd and Deeds series expansion model are shown for each case study for both resistance and reactance in Fig 1. In all studies, these plots matched very well. The mean values from the Monte Carlo simulations and the model coefficients estimate are shown in Table 1. These values agree for the first two cases but not in the third study. These discrepancies result from differences in the actual Monte Carlo simulations. Since the physical model has asymptotic tendencies that the polynomial model does not account for, the sampling has to be limited to a certain range. A more realistic choice of distribution would remedy this issue.

TABLE 2. Average values of impedance change calculated with different methods

	Surrogate MC	Model MC	Surrogate Coeffs.
Case 1	1.232-5.392j	1.232-5.393j	1.235-2.405j
Case 2	1.229-5.381j	1.229-5.384j	1.235-5.407j
Case 3	1.106-5.068j	1.160-5.063j	1.246-5.484j

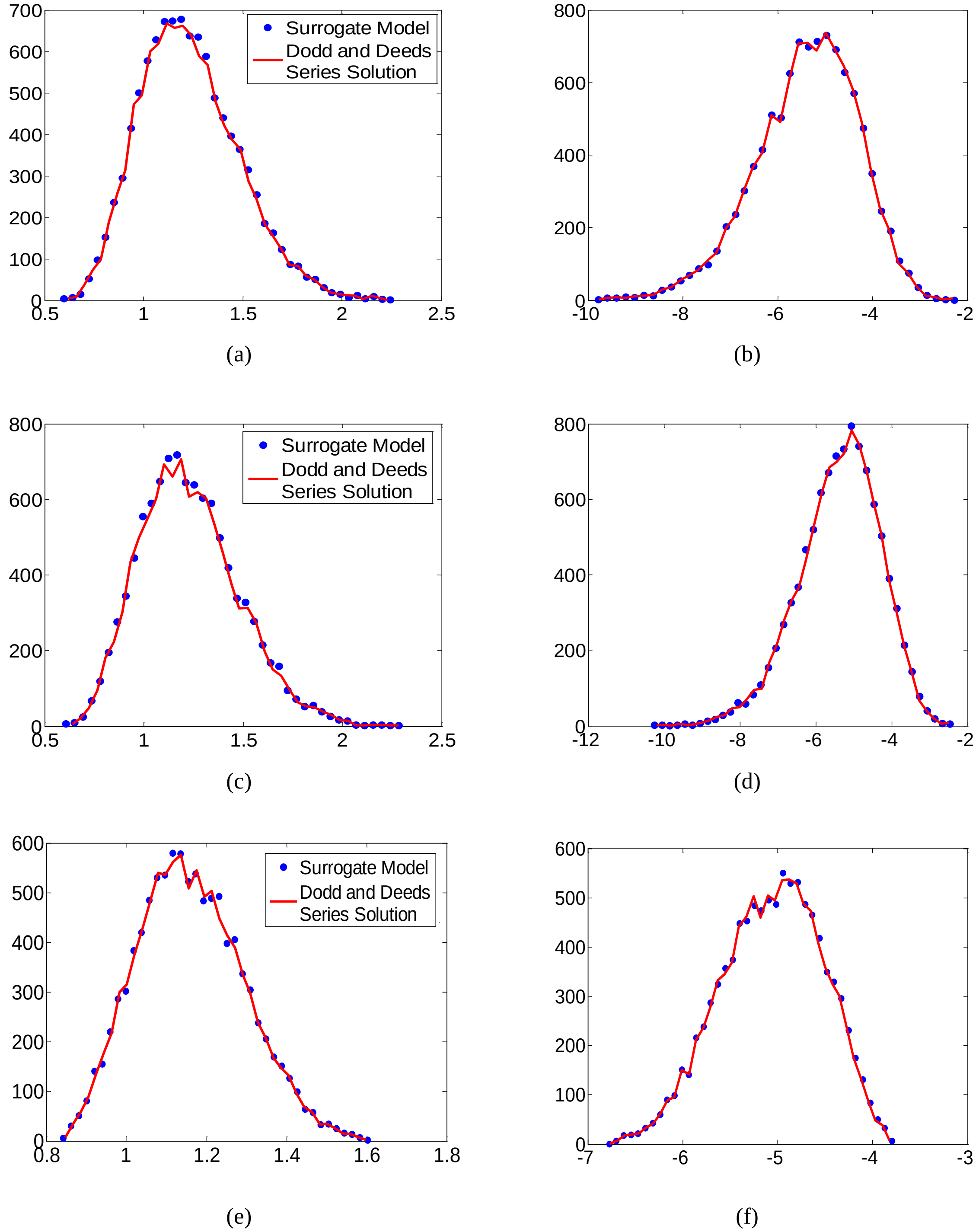

FIGURE 1. PDF of the responses for all three case studies. (a) is the PDF of the resistance from case 1 and (b) is the PDF of the reactance of case 1. (c) is the PDF of the resistance from case 2 and (d) is the PDF of the reactance of case 2. (e) is the PDF of the resistance from case 3 and (f) is the PDF of the reactance of case 3.

REFERENCES

1. J. Bowler, S. Jenkins, L. Sabbagh, H. Sabbagh, *J. of Appl. Physics* **70 (3)**, 1107-1114 (1991).
2. N. Nakagawa, T.A. Khan, J. Gray, “Eddy Current Probe Characterization for Model Input and Validation,” in *Review of Progress in QNDE 21A*, edited by D.O Thompson and D. E Chimenti, AIP Conference Proceedings vol. 509, American Institute of Physics, Melville, NY, 2000, pp. 473-480.
3. Z. Badics, Y. Matsumoto, K. Aoki, F. Nakayasu, M. Uesaka, and K. Miya, *IEEE Transactions on Magnetics* **33 (2)**, pp. 1012-1020 (1997).
4. S. K. Choi, R. Grandhi, and R. Canfield, *Reliability-based Structural Design*, Springer-Verlag, London, 2007.
5. R. Ghanem and P. Spanos, *Stochastic Finite Elements: A Spectral Approach*, Springer-Verlag, New York, 17-42, 1991.
6. W. Gautschi, *Orthogonal Polynomials: Computation and Approximation*, Oxford University Press, New York, 2004.
7. N. Wiener, *American Journal of Mathematics* **60 (4)**, 897-936 (1938).
8. A. I. Khuri, *Advanced Calculus with Application in Statistics,* Wiley-Interscience, Hoboken, N.J., c2003.
9. D. Xiu, *Numerical Methods for Stochastic Computations: A Spectral Method Approach,* Princeton University Press, Princeton NJ, 2010.
10. T. Theodoulidis, E. Kriezis, *Journal of Materials Processing Technology* **161**, 343-347 (2005).
11. A. Krommer and C. Ueberhuber, *Computational Integration,* SIAM, Philadelphia PA, c1998.
12. C.V. Dodd and W.E. Deeds, *J. of Appl. Physics* **39 (6)**, 2829-2837 (1968).
13. N. Nakagawa, T.A. Khan, J. Gray, “Eddy Current Probe Characterization for Model Input and Validation,” in *Review of Progress in QNDE 21A*, edited by D.O Thompson and D. E Chimenti, AIP Conference Proceedings vol. 509, American Institute of Physics, Melville, NY, 2000, pp. 473-480.